

WHAT IS CLAIMED IS:

1 1. In a system area network (SAN) including a source node and a
2 destination node coupled by a network fabric, with the system for transferring data
3 between a source node and a destination node, with the network fabric coupling the
4 source and destination nodes including first and second routers having multiple input
5 ports coupled to multiple output ports by a cross-bar switch, and with the SAN
6 implementing data transfers as a sequence of request/response packet pair transactions,
7 with each request and response packet containing a header including a destination field,
8 and with the SAN for implementing ordered transactions requiring that packets be
9 received in the order transmitted and unordered transactions where packets may be
10 received out of order, a system for implementing adaptive sets of lanes between said first
11 and second routers, said system comprising:

12 configuration logic at said first router for configuring an adaptive set
13 including multiple lanes, with a the control logic associating a designated input port with
14 the adaptive set and associating a unique output port with each lane of the adaptive set;
15 routing option control logic at said source node for setting adaptive control
16 bits in said destination field to specify whether the packet could use the routing
17 cappabilities of the adaptive set or should be routed down a specific lane of the adaptive
18 set;

19 routing control logic at said first router, responsive to the destination field
20 of a packet received at said designated input port, for assigning a specific output port to
21 said packet, and, if said specific output port is associated with said adaptive set,
22 adaptively assigning a port associated with a lane in the adaptive set if the adaptive
23 control bits specify adaptive routing or deterministically specifying said specific output
24 ports if said adaptive control bits specify determinist routing.

1 2. The system of claim 1 wherein:
2 said routing control logic includes a routing table with each entry in the
3 table including a bit specifying whether the entry is for an adaptive set, and if so, a field
4 identifying the adaptive set.

1 3. In a system area network (SAN) including a source node and a
2 destination node coupled by a network fabric, with the system for transferring data
3 between a source node and a destination node, with the network fabric coupling the
4 source and destination nodes including a router having multiple input ports coupled to
5 multiple output ports by a cross-bar switch, where the router may include an adaptive set
6 of lanes coupled to an input port where a designated output port is assigned to each lane
7 so that packets received at the input port may be adaptively routed on any one of the
8 multiple output ports assigned to the lanes of the adaptive set, and with the SAN
9 implementing data transfers as a sequence of request/response packet pairs, and with each
10 request packet containing a header including a destination field, a method for flushing
11 lanes in an adaptive set configured at said router, said method comprising the steps of:
12 at said source node, preparing a sequence of write packets with the
13 destination field of each packet in the sequence having adaptive control bits specifying a
14 different lane in an adaptive set;
15 at said source node, transmitting said sequence of write packets;
16 at said router, receiving said write packets, and, if an adaptive set is
17 defined, responding to the adaptive control bits of each received write packet to force said
18 packet to the output port specified by the adaptive control bits in the write packet.

1 4. The method of claim 3 further comprising the steps of:
2 at the source node, including a particular value in each of the write packets
3 and specifying a particular location at the destination node;
4 at the destination node, for each write packet, storing said particular value
5 at the specified location;
6 at the source node, accessing the particular locations at the destination
7 node and if the particular value is read from the particular locations specified by the
8 sequence of write packets indicating that the barrier transaction was successful.

1 5. The method of claim 3 further comprising the steps of:
2 at the router, limiting the number of lanes in an adaptive set to a specified
3 number;
4 at the source node, forming said selected number of write packets in said
5 sequence.

1 6. A routing topology comprising:
2 a first level including first and second first-level routers, each first-level
3 router having a first, second, and third input ports coupled to first, second, and third
4 output ports by a cross-bar switch, and with each first-level router configured to include
5 an adaptive set including first and second lanes, with the first input port associated with
6 the adaptive set and a first output port associated with the first lane and a second output
7 port associated with the second lane of the adaptive set, and with each first-level router
8 including routing logic for adaptively assigning a lane in the adaptive set to adaptively
9 route packets received at the first input port to first and second output ports associated
10 with lanes of the adaptive set;
11 a second level of routers including first and second first-level routers, each
12 second-level router having first and second input ports coupled to first and second output
13 ports by a cross-bar switch;
14 a first uplink coupling the first output port of the first first-level router to
15 the first input port of the first second-level router;
16 a second uplink coupling the second output port of the first first-level
17 router to the first input port of the second second-level router;
18 a third uplink coupling the first output port of the second first-level router
19 to the second input port of the first second-level router;
20 a fourth uplink coupling the second output port of the second first-level
21 router to the second input port of the second second-level router;
22 a source node coupled to the input port of said first first-level router; and
23 a destination node coupled to the third output port of said second first-level
24 router.

- 1 7. The routing topology of claim 6 further comprising:
2 a first downlink coupling the first output port of the first second-level
3 router to the second input port of the first first-level router;
4 a second downlink coupling the second output port of the first second-
5 level router to the second input port of the second first-level router;
6 a third downlink coupling the first output port of the second second-level
7 router to the third input port of the first first-level router; and
8 a fourth downlink coupling the second output port of the second second-
9 level router to the third input port of the second first-level router.

009260 42402960